

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

-----X

	:	11 Civ. 6188 (DLC)
	:	11 Civ. 6189 (DLC)
FEDERAL HOUSING FINANCE AGENCY,	:	11 Civ. 6190 (DLC)
	:	11 Civ. 6192 (DLC)
Plaintiff,	:	11 Civ. 6193 (DLC)
	:	11 Civ. 6195 (DLC)
-v-	:	11 Civ. 6196 (DLC)
	:	11 Civ. 6198 (DLC)
	:	11 Civ. 6200 (DLC)
JPMORGAN CHASE & CO., et al.,	:	11 Civ. 6201 (DLC)
Defendants;	:	11 Civ. 6202 (DLC)
	:	11 Civ. 6203 (DLC)
	:	11 Civ. 6739 (DLC)
And other FHFA cases.	:	11 Civ. 7010 (DLC)
	:	11 Civ. 7048 (DLC)
	:	

-----X

OPINION & ORDER

DENISE COTE, District Judge:

The above captioned actions are part of a suite of sixteen cases currently before this Court in which the Federal Housing Finance Agency ("FHFA" or "the Agency"), as conservator for Fannie Mae and Freddie Mac (together, the "Government Sponsored Enterprises" or "GSEs"), alleges misconduct on the part of the nation's largest financial institutions in connection with the offer and sale of certain mortgage-backed securities purchased by the GSEs in the period between 2005 and 2007.¹ On October 10,

¹ Also pending before this Court is FHFA v. UBS Americas, Inc., et al., 11 Civ. 5201 (DLC). The FHFA has also brought two similar actions, which are pending in federal courts in California and Connecticut. See FHFA v. Countrywide Financial Corp., et al., No. 12 Civ. 1059 (MRP) (C.D. Cal.); FHFA v. Royal

2012, FHFA filed an expert report authored by Dr. Charles D. Cowan ("the Report"), that set forth the methodology that the plaintiff intends to employ in selecting mortgage loans to be re-underwritten as part of its effort to prove its claims at trial in these cases.

This Opinion addresses a joint motion by all defendants in fifteen actions to challenge the plaintiff's sampling methodology on the basis of Federal Rule of Evidence 702 and Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993). As discussed in greater detail below, although the defendants have raised a number of issues bearing on the weight that the jury might ultimately choose to assign to the plaintiff's samples of mortgage loans and any inferences to be drawn from them, they have not demonstrated that the samples drawn through the methodology described in the Report should be deemed inadmissible. Accordingly, defendants' motion is denied.

BACKGROUND

I. Procedural History

As amended, the complaints in each of the FHFA actions assert that the Offering Documents used to market and sell

Bank of Scotland, No. 11 Civ. 1383 (AWT) (D. Conn). All capitalized terms in this Opinion have the meanings previously assigned to them in prior Opinions issued in this coordinated litigation.

Residential Mortgage-Backed Securities ("RMBS") to the GSEs during the relevant period contained material misstatements or omissions with respect to the owner-occupancy status, loan-to-value ("LTV") ratio, and underwriting standards that characterized the underlying mortgages. On the basis of these allegations, the complaints plead claims under Sections 11, 12(a)(2), and 15 of the Securities Act of 1933, 15 U.S.C. §§ 77k, 1(a)(2), of the Virginia Securities Act, VA Code Ann. § 13.1-522(A)(ii), (C); and the District of Columbia Securities Act, D.C. Code § 31-5606.05(a)(1)(B), (c). In six of the cases, the Agency has also asserted common law claims of fraud and aiding and abetting fraud against certain entity defendants.²

Following the decision on a motion to dismiss in FHFA v. UBS, see 858 F. Supp. 2d 306 (S.D.N.Y. 2012), discovery began in all sixteen of the coordinated actions. The UBS Action will be the first case to go to trial, on January 13, 2014. The next tranche of trials -- which includes FHFA v. JPMorgan Chase, 11 Civ. 6188 (DLC), and FHFA v. Merrill Lynch & Co., et al., 11 Civ. 11 Civ. 6201 (DLC) -- will begin on June 2, 2014.

² As noted in previous Opinions, the plaintiff also pleads defendants' statements regarding the credit ratings of the securities or Certificates purchased by the plaintiff as a separate category of misstatement under the Securities Act. These claims are largely derivative of the three core representations described above.

Across the sixteen cases, the plaintiff's claims implicate Certificates purchased by the GSEs (the "GSE Certificates") from 449 distinct securitizations. There is no dispute that the claims and defenses in these cases require the parties to re-underwrite at least some of the more than 1.1 million mortgage loans that back the GSEs' Certificates in order to determine, inter alia, whether the loans were originated in accordance with applicable guidelines. This process is highly labor intensive and therefore costly. The plaintiff represents that the re-underwriting of a single loan file requires at least 2-3 hours of work and costs approximately \$300-400. The defendants have not disagreed with these figures.

Recognizing the tremendous costs associated with re-underwriting more than 1.1 million loan files and the burden that production of those loan files would impose, both on the parties to these cases and on scores of third-party loan-file custodians, in December 2011, immediately after the initial pretrial conference in these cases, the Court ordered the parties to submit proposals for restricting discovery to a representative subset of loan files. On February 29, 2012, FHFA proposed that a sampling methodology be adopted that would yield a confidence interval of 95%, with a margin of error of $\pm 5\%$. The defendants in all sixteen actions submitted a joint response in which they rejected the idea of restricting discovery to a

sample of loan files and argued that "[i]f Plaintiff chooses to meet its burden of proof by relying on statistical sampling, it may submit such expert testimony, subject to challenge by Defendants under the Federal Rules of Evidence and Daubert."

Questioned about this submission at a conference on May 14, defendants represented that it had "never been [their] position in any way at any time that the only way to conduct discovery in this case or to get to the merits is by full production of every loan file." But, in a supplemental submission of June 7, defendants reiterated their objection to the use of sampling to restrict the scope of discovery, citing concerns that sampling would abridge their Seventh Amendment rights and deprive them of their ability to challenge the plaintiff's evidence pursuant to F.R.E. 702 and Daubert.

After reviewing these submissions, as well as a June 7 supplemental submission from the plaintiff, the Court announced during a conference on June 13 that it would not restrict discovery in these cases to a sample of loan files. This decision was influenced principally by the defendants' categorical refusal to consent to any sampling-based restriction of discovery and their insistence on their right at trial to reach outside of any sample to present arguments premised on other loans. Faced with defendants' position, the plaintiff reserved its right as well to respond to the defendants'

evidence taken from other loans with its own evidence from non-sample loan files.

The parties thereafter began a massive effort to gather as many as possible of the 1.1 million-plus loan files and their corresponding underwriting guidelines from scores of institutions. While some of the loan files are stored electronically, others are not. Therefore, for a substantial number of files, there may remain a question of whether the entire file has been located and reassembled.

Following the collapse of the discussions that would have restricted loan file evidence in this litigation to agreed-upon samples, the plaintiff proceeded to develop its own methodology for identifying samples of loan files to support its claims at trial. These samples will be of particular importance to the plaintiff in showing that the Offering Documents contained misrepresentations regarding the underwriting practices for the mortgages that supported each security.

The defendants requested that the plaintiff be required to give them notice, well in advance of the exchange of expert reports, of the identity of the loans in each of its samples as well as a detailed description of the deficiencies in the underwriting process that the plaintiff would be alleging for each such loan. The identification of the loans chosen by the plaintiff will permit the defendants to conduct their own re-

underwriting of the plaintiff's loan samples, and to evaluate just how well or poorly the originators, and indeed the underwriters of the securities, performed their work. As described by the defendants, the disclosure of the deficiencies that the plaintiff perceives occurred in the underwriting process will give the defendants important information that they need to better understand and respond to the loan-specific factual assertions that the plaintiff may make at trial. Such disclosures by the plaintiff would serve yet another purpose as well. They would help the defendants decide whether to present alternate sets of samples to the jury, either to serve as counter-samples to the plaintiff's samples or to support other defenses and arguments that they wish to present at trial.

The Court agreed that an exchange of information about the parties' samples in advance of the exchange of expert reports would be of assistance to all parties and helpful in the preparation of expert reports. Indeed, as the Court noted at the June 13 conference, defendants' own expert expressed the view that a sampling protocol could be settled upon and evaluated at an early stage of any litigation. Early vetting of the parties' sampling protocols is particularly important in this case, as the plaintiff and defendants should not be required to begin the costly and time-consuming process of re-

underwriting without some assurance that the samples will be deemed admissible.

To that end, on July 31 the Court directed the plaintiff to file an expert report outlining its proposed methodology for sampling in the UBS Action, 11 Civ. 5201 (DLC). The plaintiff complied with a report filed on August 9 (the "August 9 Report"). A Stipulation and Order of August 28 set a schedule for the defendants in the UBS Action to assert any challenge to that methodology pursuant to F.R.E. 702 and Daubert. The UBS Defendants declined to challenge the plaintiff's August 9 Report, thereby waiving any objection that they might fairly have asserted in response to its contents.

On October 10, the plaintiff filed the Report that is the subject of this Opinion. As outlined in greater detail below, the Report describes the plaintiff's proposed sampling methodology for the remaining fifteen cases. Pursuant to an Order of October 18, defendants' joint motion to exclude was filed on October 26 and became fully submitted on November 16. As was emphasized during the July 31 and October 13 conferences and in the October 18 Order, to the extent the present motion fails to assert any objections to the Report that might fairly have been made at the time it was filed, those arguments are waived. The present motion is being decided without prejudice to defendants' right to challenge, on Daubert grounds,

additional opinions that may be expressed by the plaintiff's expert.

Meanwhile, the parties worked on a protocol for the exchange of preliminary results from their re-underwriting.³ An Order of November 26 resolved, inter alia, remaining issues of dispute in this regard and set a schedule and procedures for the re-underwriting of loan samples in the first three actions that will proceed to trial. Pursuant to the schedule, the parties will work together to determine as soon as possible when a loan file and the associated underwriting guidelines are "the best representation" of those documents "existing at the time of the loan's origination that the parties have been able to recreate as of the time" of their agreement. The FHFA will then commence its re-underwriting of that loan. Upon completion of a re-underwriting review "for at least seventy-five (75) percent" of the loans in a particular sample, FHFA will disclose "its

³ After an initial meet-and-confer process, the parties submitted competing proposed schedules on November 5. These proposals were the subject of significant discussion at a conference on November 6. Thereafter, the parties engaged in further discussions in an effort to narrow the issues of dispute. Revised competing proposals were submitted on November 14. At a conference on November 15, after several hours of argument, the Court ruled on several key areas of dispute. On November 19, the parties submitted competing proposed stipulations memorializing the Court's oral rulings and raising additional issues in dispute.

initial factual findings" to the defendants.⁴ Within twenty-one days, the defendants must notify FHFA whether they intend to make use of any alternative set of loans in that securitization for the purposes of re-underwriting or for any other purpose, and within twenty-eight days must disclose any rebuttal findings based upon the defendants' re-underwriting of the plaintiff's sample. When the FHFA has made such disclosures for fifty percent of the securitizations in an action, the defendants must notify FHFA within twenty-one days whether they intend to make use of any set of loans drawn on any basis other than by securitization, and describe such alternative set by loan numbers.

Thus, over the course of the coming months, the FHFA will make factual assertions on a loan-by-loan basis regarding defects in the underwriting process, and the defendants will respond to those assertions and make their own disclosure of the samples on which they may rely in these litigations.

Thereafter, the parties will exchange their initial re-underwriting expert reports in these three actions, which is scheduled to occur on May 3, 2013, in the UBS Action, and on August 5, 2013, in the other two actions. Once the Court and

⁴ If the parties cannot reach agreement that they have a reasonably accurate re-creation of a loan file and corresponding guidelines for at least 75 of the 100 loans in a sample, FHFA has no obligation to make a disclosure to the defendants of its initial findings with respect to that sample.

parties have had more experience with this protocol of reciprocal disclosures it may be extended to the remaining actions, which are to be tried later, in its present or in a modified form. It is hoped that this sequenced exchange of information will assist all parties in preparing their cases for trial and in assessing the risks they face in proceeding to trial.

II. The Report and Its Author

The author of the Report, Dr. Charles Cowan, holds a Bachelors of Arts degree in Economics from the University of Michigan, a Master of Arts degree in Economics from the University of Michigan, and a doctorate in Mathematical Statistics from George Washington University. He is currently the Managing Partner of Analytic Focus, LLC, a firm that provides expert witness and consulting services in litigation, as well as statistical and econometric advice to government agencies and private financial institutions in connection with loan-portfolio analysis. Prior to joining Analytic Focus, Dr. Cowan served, at various points, as a director of financial research at PricewaterhouseCoopers, LLP; Chief Statistician for the Federal Deposit Insurance Corporation and the Resolution Trust Corporation; Chief Statistician for the National Center for Education Statistics, an agency within the Department of

Education; and Chief of the Survey Design Branch of the U.S. Bureau of the Census.

Dr. Cowan is an adjunct professor in the Department of Biostatistics in the School of Public Health at the University of Alabama. He is the author or co-author of two books and numerous articles on statistical methods and sampling, including several that pertain specifically to the use of statistical sampling or financial analysis in connection with lending institutions and loans. Dr. Cowan is a member of several professional societies and has held leadership positions in both the American Statistical Association and the American Association for Public Opinion Research.

The Report identifies 100 supporting loans randomly drawn from the supporting loan group ("SLG") or groups for each of the 427 securitizations at issue in these fifteen cases.⁵ According to the Report, these samples will enable FHFA to make estimates regarding the underlying populations with a 95% confidence level

⁵ Dr. Cowan principally relied on the loan tapes produced by the defendants to select the samples of loans. A loan tape is a collection of data concerning the individual loans in the securitization compiled by the RMBS sponsor while the securitization is being created. The tape typically includes the LTV ratio, and a borrower's credit score and occupancy status, among many other fields of data. Where the loan tape was unavailable or incomplete, Dr. Cowan used data from a recognized provider of financial data, CoreLogic.

and a margin of error of $\pm 10\%$.⁶ For example, if the Agency finds based upon its analysis of a single sample set that in 50% of cases (or 50 out of 100 files) the loan was not originated in accordance with underwriting guidelines, it could infer with 95% confidence that the defect rate for the underlying population ranges between 40% and 60%. These figures are calculated on the assumption that each of the observations of interest -- whether a loan was originated in compliance with underwriting guidelines, whether the underlying property was owner occupied, and whether the LTV of the loan was misrepresented -- can be coded in a binary, true/false basis.

The Report argues that a 95% confidence level with a maximum margin of error of $\pm 10\%$ strikes the correct balance between cost and accuracy, because increasing the sample size beyond 100 loans, even by quadrupling it, yields only marginal increases in accuracy while imposing significant re-underwriting costs. According to Dr. Cowan, as the estimated defect rate deviates from 50%, the margin of error significantly decreases. Thus, the actual margin of error in any given case will in all likelihood be smaller than $\pm 10\%$.

In an effort to further reduce the margin of error, for many of the securitizations Dr. Cowan also stratified the sample

⁶ Dr. Cowan opines that a sample size of 95 is sufficient to achieve this confidence interval, but explains that he "rounded up" to 100 loans in an abundance of caution.

by Fair Isaac Corporation ("FICO") credit score, which appears on most of the loan tapes.⁷ Dr. Cowan reports that a borrower's credit score, which can range from 350 to 850, is correlated with the likelihood that a borrower will repay a debt. The population of loans in each securitization was sorted into four roughly equal-sized groups based on whether the borrower's FICO score was very low, somewhat low, somewhat high or high by the standards of that securitization. A random number was then generated for each loan in the stratum, thereby ensuring that each loan had an equal chance of being selected. The loans in each stratum were then sorted from lowest to highest based on these random numbers. The first 25 loans in each stratum were selected for inclusion in the sample. According to the Report, this strategy will not lessen the sample's predictive power and has the potential to decrease the margin of error if the credit score is in fact correlated with one or more of the outcomes of interest.

To ensure that the sample selected for each securitization is a random and unbiased representation of the population from which it was selected, in most cases, Dr. Cowan tested the sample against the population on eleven variables drawn from the loan tapes: FICO score, debt-to-income ratio, LTV ratio, CLTV

⁷ As set forth below, stratification is a process where the population of loans is divided into mutually exclusive and exhaustive subgroups of loans.

ratio, note rate, loan amount, original term, documentation type, occupancy type, property type, and loan purpose. For continuous variables like FICO score and LTV ratio -- which have numeric values increasing or decreasing in value -- Dr. Cowan compared the mean of the sample distribution to that of the source population using a statistical method known as a z-test. For categorical variables like documentation type, a Chi-square test was used. According to the report, the results of these tests indicate "a very high level of correspondence between the samples and their populations."

The Report concludes by suggesting ways in which the results of the plaintiff's re-underwriting of sampled loans might be extrapolated to the larger population. Although the Report discusses several different extrapolation techniques that might be employed, it does not ultimately settle on one because "the determination of which method to use depends on the results of the testing on the samples." Ultimately, however, the use of sampling will assist the fact finder in determining liability not only for any one securitization with a "known level of accuracy," but also in analyzing liability for a combined set of securitizations in any one case with "a much higher level of confidence" when the estimated total number of breaches across all securitizations in one case are counted.

DISCUSSION

Federal Rule of Evidence 702 grants an expert witness testimonial latitude unavailable to other witnesses, provided that (1) "the testimony is based on sufficient facts or data," (2) "the testimony is the product of reliable principles and methods," and (3) "the expert has reliably applied the principles and methods to the facts of the case." "[T]he proponent of expert testimony has the burden of establishing by a preponderance of the evidence that the admissibility requirements of Rule 702 are satisfied." United States v. Williams, 506 F.3d 151, 160 (2d Cir. 2007). Nonetheless, the district court performs the role of "gatekeeper" -- ensuring that the proponent has made the necessary showing and that the expert's testimony "both rests on a reliable foundation and is relevant to the task at hand." Daubert, 509 U.S. at 597.

"[A] trial judge should exclude expert testimony if it is speculative or conjectural or based on assumptions that are so unrealistic and contradictory as to suggest bad faith." Zerega Ave. Realty Corp. v. Hornbeck Offshore Transp., LLC, 571 F.3d 206, 213-14 (2d Cir. 2009) (citation omitted). Moreover, in order to be admissible, "[a]n expert opinion requires some explanation as to how the expert came to his conclusion and what methodologies or evidence substantiate that conclusion." Riegel v. Medtronic, Inc., 451 F.3d 104, 127 (2d Cir. 2006), aff'd on

other grounds, 128 S. Ct. 999 (2008). An explanation is necessary because "when an expert opinion is based on data, a methodology, or studies that are simply inadequate to support the conclusions reached, Daubert and Rule 702 mandate the exclusion of that unreliable opinion testimony." Ruggiero v. Warner-Lambert Co., 424 F.3d 249, 255 (2d Cir. 2005) (citation omitted).

While a district court has "broad latitude" in deciding both "how to determine reliability" and in reaching "its ultimate reliability determination," it may not abandon this "gatekeeping function." Williams, 506 F.3d at 160-61 (citation omitted). "[N]othing in either Daubert or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the ipse dixit of the expert." Kumho Tire Co. v. Carmichael, 526 U.S. 137, 157 (1999) (citation omitted).

In seeking to exclude the plaintiff's samples, defendants do not challenge Dr. Cowan's qualifications or the scientific reliability of statistical sampling generally. Rather, they identify numerous "methodological errors" that they argue, taken together, render the Report unreliable and therefore make the samples and inferences to be drawn from the analysis of the samples inadmissible. These issues will be addressed in turn.

I. Failure to Produce Underlying Data

Defendants' first argument is that FHFA has failed to produce sufficient data to permit their own expert to determine that the samples chosen in the October 10 Report are representative, random, unbiased, and not subject to manipulation. Specifically, they complain that they were not provided access to:

(1) what loan data FHFA provided to Dr. Cowan; (2) the full set of calculations, computer programs, computer spreadsheets, and data files generated or used by Dr. Cowan that connect the loan population to the selected sample; (3) the tests of sample representativeness conducted by Dr. Cowan and referenced in the Report; (4) documentation concerning the methodology by which Dr. Cowan generated allegedly random numbers for his sampling methodology; and (5) [information] whether multiple samples from each securitization were drawn and tested.

This information is largely irrelevant to the task of assessing the scientific validity of Dr. Cowan's sampling protocol and the admissibility of the samples drawn through use of that protocol. In any event, the defendants have been provided with most of this information. With respect to loan data, the Report (as well as some additional information provided to the defendants) identifies by Bates number the documents on which Dr. Cowan relied to draw his sample. As explained in the Report, in certain instances, these materials, which were produced during party discovery, were supplemented using information obtained from CoreLogic, a commercial provider

of financial data. Dr. Cowan has also affirmed that he provided defendants' expert with "the computer programs [he] used . . . which include the starting point for the fixed sequence of random numbers known as the 'seed.'" As for the tests of sample representativeness, the Report relates p-values -- which quantify the likelihood that the sample was representative -- for each of the tests that Dr. Cowan conducted on the samples using standard statistical methods.⁸ And defendants' objection that the Report fails to indicate whether multiple samples were tested -- an issue that is not mentioned by their expert -- is not relevant to assessing whether the methodology used to draw the samples is sound. If it is, any resulting sample will have the same confidence interval and margin of error, whether drawn on the first try or the fiftieth.

Defendants' real concern is that they have been unable to replicate fully the results set out in the Report, which they argue raises questions about its reliability. Some of the obstacles have been practical. FHFA acknowledges that, due to a misunderstanding regarding the composition of the SLGs, an initial version of the Report identified samples that, in 25 of the 449 sample selections, erroneously included non-supporting

⁸ The higher the p-value, the stronger the agreement between the sample and the population from which the sample is drawn. This is a standard measure of how representative the sample is of the population.

loans. That misunderstanding has now been corrected, and although defendants continue to press the argument that the initial error raises questions about the reliability of Dr. Cowan's sampling methodology, they have identified no other instances in which samples were drawn from an over-inclusive population. Defendants also complain that they have been unable to match fully the loan identification numbers "used internally by Defendants" with those used by the plaintiff in identifying its samples, but they do not contest the FHFA's contention that the information necessary to do this matching is in their possession.

In one instance, however, the defendants' inability to replicate the results expressed in the Report may arise from an error by the plaintiff's expert. The defendants' expert claims to have been unable to replicate the results of the representativeness tests expressed in Appendix 2 to the Report using the methodology described therein.⁹ Replicability is, of course, an important test in determining scientific reliability. But the representativeness checks that defendants have not been

⁹ Defendants' expert also asserts that for two of the securitizations, "the computer code provided by Dr. Cowan in support of his random selection of the samples does not replicate his sample draws." But given that the samples were purportedly drawn at random, neither he nor the defendants in their briefs provide sufficient detail regarding the nature of the replication problem. In any case, any defect in this regard would bear only on the admissibility of the two samples at issue, not the plaintiff's overall methodology.

able to replicate are not part of the sample-selection methodology; they are instead themselves tests of that methodology as applied in specific instances. Indeed, in most of the cases cited by the defendants, their own representativeness checks suggest a higher degree of certainty that the samples accurately reflect the characteristics of the population as a whole than the checks performed by the plaintiff. Moreover, the differences between defendants' own representativeness results and those in the Report are, in most cases, relatively minor.

At worst, defendants' inability to replicate exactly the plaintiff's representativeness results raises questions about the extent to which Dr. Cowan has carefully and consistently applied (or reported) his own protocol for developing and testing his samples. But these concerns are best addressed in assigning evidentiary weight to the samples and any conclusion that the plaintiff may draw from them; they do not warrant rejection of the sampling methodology or samples altogether.

II. Differences Among Certificates and Originators

Next, defendants argue that the plaintiff's sampling methodology is flawed because it will not permit the fact finder to draw conclusions on a certificate-by-certificate and originator-by-originator basis. This argument turns on two

assumptions regarding what facts the plaintiff will be required to prove at trial in order to succeed on its claims.

With respect to the first assumption, defendants are correct that "[e]ach certificate is a separate security," and as such, to the extent the plaintiff seeks to recover on the basis of the GSEs' purchase of multiple Certificates in a securitization, it must show a misstatement of fact that was material in the context of each transaction. See 15 U.S.C. § 77k. The plaintiff does not dispute the requirement that it prove its case for each of the Certificates on which it seeks to recover. It argues, however, that in the case of the 37 securitizations from which the GSEs purchased multiple Certificates backed by different SLGs, the Prospectus Supplements "contained misrepresentations regarding underwriting guidelines generally applicable to all classes of securities" included in the offering. (Emphasis in the original.)¹⁰ Defendants deny that this is so. The issue will thus turn on the precise language of each particular Prospectus Supplement. If the plaintiff is correct that each of the Prospectus Supplements contained such global representations, proving their falsity would be sufficient to prove a misstatement material to the GSEs' purchase of Certificates included in those offerings.

¹⁰ In 412 securitizations, the GSEs purchased one or more Certificates backed by a single SLG.

Cf. NECA-IBEW Health & Welfare Fund v. Goldman Sachs & Co., 693 F.3d 145, 163 (2d Cir. 2012) (noting that, whether in the Shelf Registration Statement or in the Prospectus Supplement, "the location of the representation has no effect on a given purchaser's assertion that the representation was misleading").

The plaintiff also argues, however, that the results of its sampling will allow it to make predictions regarding underwriting breaches on a certificate-by-certificate basis. As defendants note, because only a portion of the plaintiff's samples in multiple certificate cases will be drawn from the SLGs for each of the Certificates at issue, the power of those 100-loan samples to make predictions regarding the discrete SLGs will be lower than 100-loan samples drawn from single-Certificate securitizations and likely lower than the plaintiff's target of a 95% confidence level with a margin of error $\pm 10\%$. Defendants cite the example of the CSFB 2005-12 Securitization, for which the plaintiff's sample includes 28 loans drawn from SLG 2, 46 from SLG 4, and 26 from SLG 5, with each group backing a different Certificate. They calculate that for an estimated breach rate of 50%, the group-by-group predictions based on the sample will be subject to margins of error of $\pm 18.4\%$, $\pm 14.3\%$, and $\pm 19.2\%$, respectively.

Of course, the plaintiff's sample will enable it to predict with 95% confidence the breach rate for the three SLGs combined

within a margin of error of $\pm 10\%$. FHFA suggests in its opposition brief that this level of accuracy, combined with the fact that (1) the sponsor, depositor and underwriters were the same for all Certificates offered in a single securitization, and (2) all loans included in a single securitization were generally of the same vintage and largely originated by the same entity, will permit the jury to make a finding with regard to the falsehood of Certificate-specific statements. That may or may not be so.¹¹ In choosing to use a blunter sampling instrument with respect to securitizations with Certificates backed by different SLGs, the plaintiff runs the risk that its proof will be found wanting.¹² But that is not an issue to be decided in the context of this Daubert motion. Rather, a jury will decide what weight to assign the plaintiff's samples after considering arguments that the defendants will no doubt make regarding the inadequacy of those samples and the plaintiff's rebuttals.

¹¹ In their reply, the defendants point out that in one securitization in which the plaintiff purchased multiple Certificates backed by different SLGs, each SLG contained loans originated by a different entity and the Offering Documents contained different representations regarding LTV ratios and owner-occupancy rates.

¹² It appears that modification of the plaintiff's sampling protocol to designate a separate sample for each of the 488 SLGs would require it to underwrite 3,900 additional loans across all sixteen cases.

Moreover, at the trial the jury may be able to draw inferences across securitizations based on samples drawn from many securitizations with shared characteristics, such as the identical originators, sponsors or underwriters. For instance, in the Chase Action, 11 Civ. 6188, there are 103 securitizations at issue, only 13 of which have multiple Certificates backed by separate SLGs. Thus, by the time of trial, any apparent need to provide a 100-loan sample for each of the separate SLGs supporting different Certificates within a securitization may evaporate.

Defendants also fault the plaintiff's sampling technique for failing to ensure that the plaintiff will be able to make statistical predictions about underwriting breaches on an originator-by-originator basis. But as has been previously emphasized,

FHFA's claim here is not that the originators failed to scrutinize loan applicants adequately in general; it is that defendants failed to act diligently to ensure that, consistent with the representations in the offering materials, the originators' questionable practices did not lead to the inclusion of non-conforming loans in the particular securitizations sold to the GSEs.

UBS I, 858 F. Supp. 2d at 321. Accordingly, the Agency need only show that a significant number of the loans included in the securitizations were not originated in accordance with the applicable guidelines. Differences in the guidelines that were

applied by different originators can be accounted for in coding the samples. Indeed, the November 26 Expert Scheduling Order requires the parties to account for such variation in their re-underwriting by seeking agreement on the precise guidelines that apply to a particular loan before attempting to identify instances of breach.

III. Remaining Arguments

Like their argument regarding the replicability of Dr. Cowan's significance tests, many of the defendants' remaining arguments seek to shoehorn into the Daubert rubric considerations that properly go to the weight of the plaintiff's proof. They argue, for example, that the Report's assumption that underwriting breaches can be expressed in a binary fashion may result in under-sampling because of variations in individual borrowers, underwriting criteria, and standards for exceptions. But, as already noted, the plaintiff intends to account for such variation in the coding process.

Likewise, defendants' argument that the Report unjustifiably expands the margin of error for the sampling protocol is an argument about the persuasive power of the plaintiff's proposed sample, not its admissibility. Defendants make much of the fact that Dr. Cowan previously proposed a sampling protocol that would result in a margin of error of $\pm 5\%$. But that proposal was made in support of the plaintiff's

proposal to restrict discovery in these cases to samples to which all parties would be confined in supporting their claims or defenses. Defendants' rejection of that proposal has required the parties to obtain over 1,100,000 loan files and introduced the possibility that the parties may rely on competing samples or "Alternative Sets," thereby greatly increasing the potential re-underwriting obligations for all concerned. The plaintiff's determination that a margin of error of $\pm 10\%$ is adequate to prove its case may well have been influenced by the anticipation of this additional re-underwriting burden, as well as its expectation that the samples for each of the multiple securitizations at issue in any single action will provide significant corroboration for the results shown by any single sample.

Nor does the plaintiff's decision to stratify its loan samples by FICO score render the sampling protocol unreliable and therefore worthy of exclusion. Although defendants' expert questions whether stratification by FICO score will decrease the margin of error, he does not take issue with the plaintiff's assertion that the decision to stratify will not expand the margin of error beyond the target rate of $\pm 10\%$ that the plaintiff used in choosing its sample size.

Finally, defendants fault the Report for failing to choose an extrapolation methodology and failing to specify a

substitution procedure for loans included in the sample whose files are missing or unobtainable. Neither of these objections is a valid basis for exclusion of the sample.

With respect to substitution, Dr. Cowan notes in a footnote to the Report that "if a material number of loans do not have associated loan files, [he] may draw a supplemental sample." Defendants express concern that missing files may be correlated -- occurring, for example, more commonly in paid-off loans -- and thus that the supplemented sample may not be fully representative of the population. But as Dr. Cowan notes, the need for substitution is common in research and is not a reason to reject a sampling approach altogether. In any case, the defendants have been given notice of the plaintiff's initial draws and will thus be able, after their own investigation, to raise any such arguments in disputing the weight to be afforded to the plaintiff's final samples.

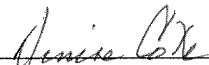
As for extrapolation, as noted, the Report suggests several potential approaches but declines to settle on one, noting that "the determination of which method to use depends on the results of the testing on the samples." To the extent the plaintiff's expert has not yet opined definitively on any issue, defendants' arguments are premature.

CONCLUSION

Defendants' October 26 motion to exclude is denied.

SO ORDERED:

Dated: New York, New York
December 3, 2012



DENISE COTE
United States District Judge